

# Price Drivers in Prediction Markets: An Agent-Based Model of Competing Narratives

Arwa Bokhari<sup>1,2</sup> 

<sup>1</sup>Department of Computer Science, University of Bristol, Bristol BS8 1UB, U.K.

<sup>2</sup>Department of Information Technology, College of Computers and Information Technology, Taif University, Taif, 21944, Saudi Arabia

**Keywords:** Narrative Economics, Opinion Dynamics, Reinforcement Learning, ABM Calibration, Prediction Markets.

**Abstract:** In this paper, I investigate price formation in prediction markets via an agent-based model (ABM). Prediction market prices can be interpreted as the probability of an event occurring, based on the aggregated beliefs of market participants. By utilizing a simple market exchange populated with opinionated agents and calibrating the model parameters, I aim to identify the effect on market price introduced by the three main drivers of the opinion formation process within two competing groups of agents: self-reinforcement; herding; and additive responses to inputs. Using a real-world dataset of Bitcoin prices, I show that both groups tend to follow the overall market sentiment. However, when the market mood aligns with a particular group's opinion, that group becomes more self-reinforcing; conversely, when the mood does not favour their opinion, they become less self-reinforcing. Furthermore, I propose to use the temporally generated parameter values—produced by the calibrated model—as well as the temporal prices and market moods shifted by seven days as the training set for a supervised machine learning and solve the multi-target learning problem to forecast both short-term price trends and the expected trajectory of the two groups' opinion dynamics. The code from this research is available for other researchers to use, build upon, and extend.

## 1 INTRODUCTION

In academic literature, terms such as *information markets*, *decision markets*, and *forecasting markets* are often used interchangeably to refer to prediction markets. (Berg and Rietz, 2003) define prediction markets as markets that primarily aim to aggregate information in order to forecast future events. These markets can also serve as decision support systems, providing insights into current situations or being used to evaluate decision-making processes. Although stock markets and prediction markets share similarities, they differ in their primary purposes. Stock markets focus on resource allocation, risk trading, and capital raising, with information aggregation being a secondary feature. Prediction markets, however, are specifically designed to aggregate information. Additionally, contracts in stock markets are based on the value of real assets, whereas prediction market contracts are linked to event outcomes and have no intrinsic value. Digital currency markets, in this context, are relatively similar to prediction mar-

kets, as both are based on non-intrinsic values.

Prediction markets are often used as a means of leveraging collective intelligence, or the “wisdom of the crowd”. Over the past decade, these markets have shown a remarkable degree of accuracy, demonstrating through various statistical tests that they outperform professional forecasters and polls (Luckner et al., 2011). Nevertheless, they are subject to biases and manipulation, which can negatively impact the corresponding financial markets (Restocchi et al., 2023). In order to overcome these challenges, there is a need for models that can accurately reproduce the dynamics of price and opinion formations. These models are required to be simple enough to be understandable and tractable while complex enough to allow for creating the realism of market dynamics.

In prediction markets, traders exchange contracts with prices that depend on the outcome of a future event. A rational buyer will purchase a contract if they believe it is undervalued and sell it if they believe it is overvalued. Until the actual outcome of the future event is revealed, traded prices reflect the collective opinions and beliefs of market participants regarding

 <https://orcid.org/0000-0003-2987-4601>

the likelihood of the event's occurrence.

More generally, in an efficient prediction market, the market price reflects all available information (Luckner et al., 2011). Thus, the evolution of opinions can be likened to the process of price formation. By employing opinion dynamics within the framework of a prediction market, we can simulate opinion evolution and how that is reflected in price movements.

With the expansion of the popularity of social networks and the rapid increase in information dissemination in this digital era, it is evident that communicated narratives do significantly influence people's opinions. This impact is especially crucial in prediction markets, where participants make forecasts on events—ranging from elections to economic trends—based on their formed opinions. Prediction markets rely on the collective wisdom of participants, who analyze data and apply their insights to predict outcomes. Therefore, it is essential to understand the nature of narratives circulating in digital media and how they shape the interpretation of available information. In many cases, two or more competing narratives about the same event coexist, and each can vary in their dominance and strength during the event under prediction. For example, during national general elections, one narrative might emphasize one candidate's strengths and their positive track record, while another competing narrative embraces other candidates' eligibility to win. These narratives often create polarized opinion-formed groups, each aligning with the story that echoes their beliefs, values, or expectations. Each group can be stronger at some periods and weaker in others, this illustrates how, in reality, different groups can interpret the same event or piece of information through entirely distinct lenses and that the collective opinions they are forming are rich, complex and nonlinear. Modeling these groups' dynamics and understanding how they interact and influence each other is crucial for identifying the underlying drivers of the opinion formation process.

In this study, I analyze three drivers of group dynamics, influenced by (Leonard et al., 2021): (1) **Self-reinforcing dynamics**, which refers to the group behaviour when it supports its own opinions and communicates mostly among itself; (2) **Herding behaviour**, where one group follows the others, often leading to convergence to one particular opinion; and (3) **Additive response to inputs** represents the situation when a group receives repetitive external inputs, assuming that each input contributes to the overall formation of opinions.

This study aims to enhance our understanding of how each of the aforementioned dynamics contributes to the process of opinion formation within each group,

the interaction dynamics between the two groups, and which dynamic most accurately influences the market price fluctuations over time as observed in real financial markets. By comparing the influence of each dynamic on opinion formation, I aim to identify the optimal parameter setting that best replicates observed patterns in real-world market data.

Parameter calibration is an essential process for the validation of ABMs because it allows the simulated model to be fitted to real-world phenomena. Without proper calibration, it is difficult to trust the simulated behavior. By tuning the model's parameters based on empirical data or observed system behavior, we can improve the alignment between the model and the actual system it is intended to represent (Song et al., 2021). Reinforcement learning has been used in ABM calibration (see (Glielmo et al., 2023)), where it is applied to calibrate the model's parameters based on feedback from the system's performance compared to the real-world counterpart. The entire model iteratively refines its parameters to better align with real-world data, optimizing the simulation's accuracy over time.

The novelty of this paper is my introduction of a new ABM for prediction markets for which I demonstrate parameter calibration, after which the model accurately fits real-world cryptocurrency price and sentiment data. Thus, allows for the following contributions (1) identify the key drivers of the opinion dynamics evolution within two groups of competing opinions, (2) build a machine-learning dataset and train a machine-learning model to predict the short-term price movement and the corresponding group dynamics between two groups of agents each of which holds an opposing opinion.

## 2 MODEL

In this section, I describe the models governing the temporal dynamics of opinion evolution and the corresponding price formation. I adopt the opinion dynamics model from (Bizyaeva et al., 2020) and the simple prediction market model from (Restocchi et al., 2023). In this context, the predicted event represents whether the market price is expected to rise or fall.

### 2.1 BFL Opinion Dynamics Model

Let  $N$  be the number of agents in the market. The agents are connected in an undirected uniformly weighted network. Agents in the network belong to one of two groups, one with a positive opinion  $N_p$

and the other  $N_n$  with a negative opinion such that  $N_p + N_n = N$ . Each agent is characterized by a scalar opinion  $x \in [-1, 1] \in \mathbb{R}$ . An agent  $i \in N_p$  updates her opinion according to the following dynamics from (Bizyaeva et al., 2022).

$$\tau_x \dot{x}_i = -x_i + \hat{S}(\alpha_p \hat{x}_p + \sigma \hat{x}_n) + b_p \quad (1)$$

Similarly, an agent  $j \in N_n$  updates her opinion according to the following dynamics from (Leonard et al., 2021):

$$\tau_x \dot{x}_j = -x_j + \hat{S}(\alpha_n \hat{x}_n + \sigma \hat{x}_p) + b_n \quad (2)$$

where  $\tau_x$  is a time scale,  $\hat{x} = \frac{1}{N} \sum_{i \in I_p} x_i$  is the group's average opinion, and  $\hat{S}$  is a saturation function defined as

$$\hat{S}(x) = \frac{1}{2} (S(x) - S(-x)) \quad (3)$$

Here,  $S$  is an odd sigmoid function. Following (Leonard et al., 2021), I use the hyperbolic tangent function, “tanh”, where  $\tanh(-x) = -\tanh(x)$ , giving  $S_z(x) = \tanh(x)$ . The parameters  $\alpha_p$  and  $\alpha_n$  are the self-reinforcing parameters, while  $b_p > 0$  and  $b_n < 0$  are the additive input parameters. These parameters are considered *positive feedback* parameters, as large positive values of  $\alpha_p$  and  $\alpha_n$  lead to strongly positive opinions ( $\alpha_p x_i > 0$ ) and strongly negative opinions ( $\alpha_n x_j < 0$ ), respectively. Additionally, a significant magnitude of  $b_p$  and  $b_n$  further drives  $x_i$  more positively and  $x_j$  more negatively, respectively. Finally,  $\sigma \in \{-1, 1\}$  represents the *herding* parameter: when  $\sigma = -1$ , the two groups are driven into a state of disagreement (anti-herding), whereas when  $\sigma = 1$ , the groups are driven into agreement (herding).

## 2.2 Prediction Market Model

In prediction markets, traders exchange contracts at prices based on the probability of the outcome of an event. If the event occurs, the contract's payoff is 1; otherwise, it is zero. Borrowing the justification and definitions from (Restocchi et al., 2023), the price in a completely efficient prediction market is defined as  $\pi = P(\text{event occurs})$ , representing the true probability of that event occurring. Hence, since  $x \in [-1, 1]$  we need normalization to constrain opinions between zero and one. Assuming traders are rational, a trader will buy if she believes the asset is undervalued, meaning her opinion is less than the market price, i.e.,  $x < \pi$ . Conversely, she will sell if she believes the asset is overvalued, i.e.,  $x > \pi$ , and will neither buy nor sell when she believes the price reflects true value, i.e.,  $x = \pi$ .

Following (Restocchi et al., 2023), market price is driven by excess demand, where trader  $i$ 's demand at

time  $t$  is defined as the difference between her opinion at time  $t$  and the current market price:

$$D_i(t) = x_i(t) - \pi(t) \quad (4)$$

Thus, the more agents believe the asset is undervalued, the higher the demand. Formally, the excess demand at time  $t$  is given by

$$ED(t) = |\vee| \sum_i D_i(t) \quad (5)$$

where  $\vee \sim N(0, 0.05)$  (see (Restocchi et al., 2023) for more details). In each trading time-step, the market price is updated according to the excess demand, with trimming to keep it between zero and one.

## 2.3 Price Drivers

Recall the dynamics from (1) and (2), and if we can observe the market and determine the market mood, then following the methodology from (Leonard et al., 2021), we can systematically identify the price drivers, assuming the prices are directly linked to opinions as I discussed previously. Starting by identifying a market mood indicator for each of the trading groups, as outlined in (Leonard et al., 2021),

$$I_p(t) = f(MM(t) + I_0), \quad (6)$$

$$I_n(t) = f(-MM(t) - I_0), \quad (7)$$

where  $f$  is a function such that  $f(0) = 0$  and  $I_0 > 0$  is the basal opinion drive.

In general, traders tend to focus more on significant shifts in market sentiment rather than minor fluctuations<sup>1</sup>. This is evidenced by dramatic market movements following events such as elections<sup>2</sup>. To accurately account for this noticeable human behaviour, the model includes a “dead zone”. We can apply a threshold region where insignificant market mood fluctuations remain ineffective in persuading agents' opinions. By introducing the conceptual “dead zone”. The function  $f$  is defined as a nonlinear function proposed by (Leonard et al., 2021) given by:

$$f(x; U, L) = \begin{cases} x - U, & \text{if } x \geq U \\ 0, & \text{if } -L < x < U \\ x + L, & \text{if } x \leq -L \end{cases} \quad (8)$$

Where  $U$  and  $L$  are the upper and lower sensitivity thresholds, both of which are non-negative. To account for the different drivers, we can set  $U_\alpha$  and  $L_\alpha$  as sensitivity thresholds for self-reinforcement, and  $U_b$  and  $L_b$  for additive input. Then, these two indicators can be used to control the dynamics of the three drivers as follows:

<sup>1</sup>See Reuters, Traders chase post-election stock gains in US options market, November 15, 2024.

<sup>2</sup>See News.com.au, Bitcoin price hits record high as Donald Trump moves closer to victory.

### 2.3.1 Self-Reinforcement

The rates of change for the reinforcement parameters  $\alpha_p$  and  $\alpha_n$  are directly proportional to market mood inputs, with a common proportionality constant  $k_\alpha$ , as described in (Leonard et al., 2021). This proportional relationship implies that shifts in market sentiment drive the dynamics of  $\alpha_p$  and  $\alpha_n$ , scaling their strength according to  $k_\alpha$ . Specifically, the differential equations governing these rates of change are:  $\dot{\alpha}_p = k_\alpha I_p(t)$  and  $\dot{\alpha}_n = k_\alpha I_n(t)$ .

In these equations,  $I_p(t)$  and  $I_n(t)$  represent the mood inputs associated with positive and negative market mood, respectively. The parameter  $k_\alpha$  acts as a sensitivity factor, determining how strongly  $\alpha_p$  and  $\alpha_n$  respond to changes in  $I_p(t)$  and  $I_n(t)$ . As market sentiment fluctuates, these dynamics enable  $\alpha_p$  and  $\alpha_n$  to adjust proportionally, reflecting the groups' response to the current mood of the market.

### 2.3.2 Herding

The herding parameter,  $\sigma(t)$ , switches based on the market mood at time  $t$  and is defined as follows:

$$\sigma(t) = \begin{cases} 1 & \text{if } MM(t) \in \{-1, +1\} \\ -1 & \text{otherwise} \end{cases}$$

This definition means that when the market mood is extreme (either  $+1$  for strongly positive or  $-1$  for strongly negative), the groups will be herding, controlled by  $\sigma(t) = 1$ . Conversely, when the market mood does not reach these extreme values,  $\sigma(t)$  is set to  $-1$  allowing the groups to act independently or in opposition.

### 2.3.3 Additive Inputs

The dynamics of the additive input parameters  $b_p$  and  $b_n$  are governed by the following differential equations, exactly analogous to  $\dot{\alpha}_p$  and  $\dot{\alpha}_n$ , discussed previously.  $\dot{b}_p = K_b I_p(t)$  and  $\dot{b}_n = K_b I_n(t)$ .

## 3 RESULTS

The dataset used in this study is sourced from a public repository<sup>3</sup>. It includes daily records of market sentiment derived from Yahoo Finance and Alternative.me, alongside the closing prices of Bitcoin, covering the period from February 1, 2018, to March 31,

<sup>3</sup><https://www.kaggle.com/datasets/adilbhatti/bitcoin-and-fear-and-greed>

2023. This dataset can be used in the context of prediction markets as it captures both price fluctuations and market mood, providing insights into how sentiment influences cryptocurrency trading decisions. Since Bitcoin trading is highly sentiment-driven and the market operates without a central authority, the dataset's sentiment scores and price data reflect the underlying psychological factors that are often central in prediction market dynamics.<sup>4</sup>

### 3.1 Parameter Calibration

The parameter combinations to be calibrated are defined as follows:  $U_\alpha$ ,  $L_\alpha$ ,  $K_\alpha$ , and  $P0_\alpha$  each have a lower bound of 0.1, an upper bound of 1.0, and a precision of 0.1. The parameter  $\sigma$  ranges from -1 to 1.0 with a precision of 2. Similarly,  $U_b$ ,  $L_b$ ,  $K_b$ , and  $P0_b$  have lower and upper bounds of 0.1 and 1.0, respectively, with a precision of 0.1. These ranges and precisions result in a total of two hundred million possible parameter combinations for calibration.

Formulating the calibration process as an optimization problem, it is evident that an exhaustive search for the optimal parameter setting with the lowest loss is infeasible. Therefore, I apply a search method that uses reinforcement learning proposed by (Glielmo et al., 2023), which utilizes a simple  $\epsilon$ -greedy algorithm to balance exploration and exploitation in parameter selection, with a fixed learning rate. The reinforcement learning algorithm uses an ensemble of three samplers—Random Forest, XGBoost, and Best Batch—to explore the parameter space. Each sampler selects a sample, runs the agent-based model (ABM) simulation with the chosen parameters, and applies a loss function to evaluate the fitness of the simulated results compared to the real-world time series. This iterative process continues until either the maximum number of search rounds 10000 is reached or convergence is achieved. The output of the calibration process is the parameter vector associated with the lowest loss. In this experiment, I applied the Euclidean distance loss function. The optimal parameters, associated with the minimum loss, resulted in a loss of 0.15 and the following calibrated values:  $U_\alpha = 0.2$ ,  $L_\alpha = 0.8$ ,  $K_\alpha = 0.3$ ,  $P0_\alpha = 0.7$ ,  $\sigma = 1.0$ ,  $U_b = 0.3$ ,  $L_b = 0.1$ ,  $K_b = 0.8$ , and  $P0_b = 0.1$ .

<sup>4</sup>Note, here I use the ready-to-use software from (Benedetti et al., 2022) for calibration, and I use the prediction market exchange simulation from (Restocchi et al., 2023).

### 3.2 Model Evaluation

As a baseline, I am using the market mood as a direct representation of traders’ opinions, assuming that the aggregated sentiment—scaled from  $-1$  (negative) to  $+1$  (positive)—reflects the collective expectations and beliefs of traders regarding Bitcoin’s future price movements. By treating market mood as a proxy for individual trader opinions, we can model opinion dynamics in a simplified manner, where shifts in market sentiment directly influence trading behaviour. This baseline provides a straightforward foundation to assess the impact of sentiment on price, against which I will compare the calibrated model.

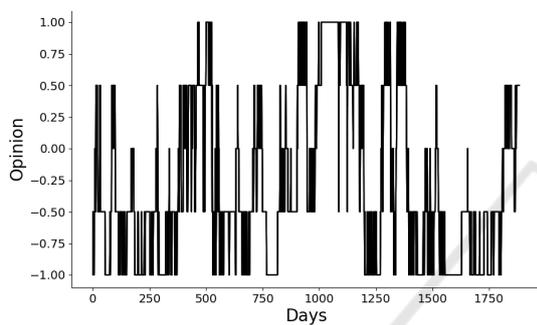


Figure 1: Uniform group opinion dynamics using Market Mood as the opinion.

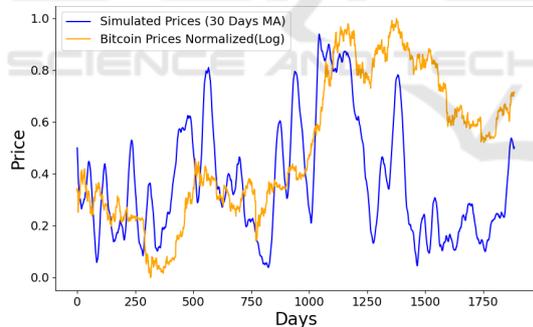


Figure 2: The corresponding price dynamics from the opinion dynamics in Figure 1.

Figure 1 illustrates the uniform group opinion dynamics when market mood is used as the opinion. The figure shows how opinion levels fluctuate over time based only on market mood<sup>5</sup>. Figure 2 depicts the corresponding price dynamics resulting from the opinion dynamics shown in Figure 1. This figure demonstrates how variations in opinions, driven by market mood, impact simulated prices over time. As market mood fluctuates, these shifts in collective

<sup>5</sup>Hence, no opinion dynamics model is used here, as the market mood is treated as a uniform opinion, with all agents sharing the same opinion at each time step.

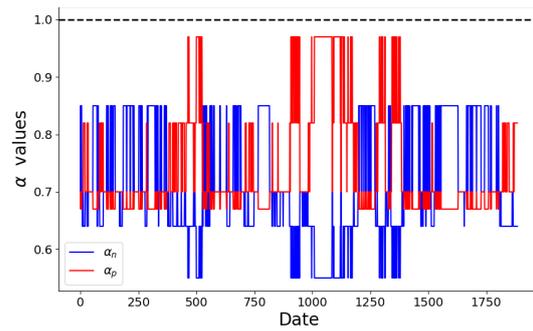


Figure 3: Alpha dynamics applying  $U_\alpha = 0.2$ ,  $L_\alpha = 0.8$ ,  $K_\alpha = 0.3$ ,  $P0_\alpha = 0.7$ . This figure illustrates the dynamics of two self-reinforcing parameters,  $\alpha_p$  (in red) and  $\alpha_n$  (in blue), over time. These parameters represent the sensitivity or reinforcement levels for the Positive Group and Negative Group, respectively, as they respond to market mood.

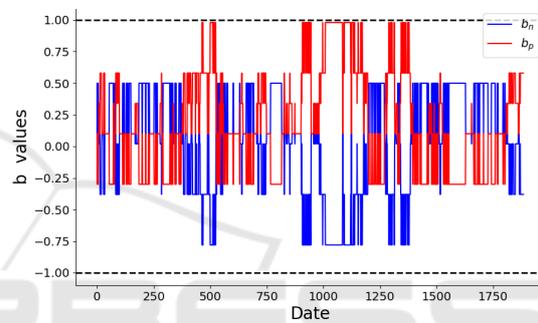


Figure 4: Input dynamics applying  $U_b = 0.3$ ,  $L_b = 0.1$ ,  $K_b = 0.8$ , and  $P0_b = 0.1$ . This figure shows the dynamics of the input parameters  $b_p$  (in red) and  $b_n$  (in blue) over time, representing external input factors that influence the positive group and negative group, respectively.

sentiment are reflected in price movements.

Figure 3 illustrates the dynamics of the self-reinforcing parameters  $\alpha_p$  and  $\alpha_n$  over time, corresponding to the self-reinforcement of the positive and negative groups, respectively. When the market mood is positive,  $\alpha_p$  tends to increase, which means the positive group is more self-reinforcing. Conversely, when the market mood is negative,  $\alpha_n$  tends to rise, meaning that the negative group is more self-reinforcing. This pattern shows that the values of  $\alpha$  adapt to prevailing market sentiment, altering the reinforcement of each group’s opinion based on the overall market mood. However, recall the dynamics (1) and (2) and since,  $\sigma = 1$ , then the group’s self-reinforcing parameter  $\alpha$  needs to exceed the critical value, i.e.  $\alpha > 1$  for the group to allow for the positive feedback  $\alpha\hat{x}$  to dominate. As shown in Figure 3  $\alpha$  values never reach 1, both groups will be more influenced by each other’s opinions. Hence,  $\sigma$  controls the tendency to herding.

Figure 4, shows that when the market mood is positive, the magnitude of input parameter  $b_p$  for the pos-

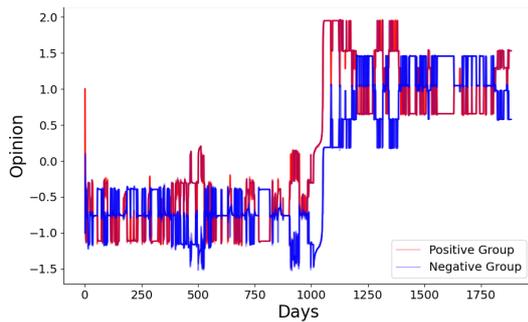


Figure 5: Opinion evolution for two distinct groups in the market: a “positive group” (represented in red) and a “negative group” (in blue). Group opinion dynamics when applying the optimal parameters setting.

itive group tends to be higher, adding to the group’s positive opinion. This positive external input aligns with the market sentiment, strengthening the positive group’s opinion. Conversely, when the market mood is negative, the magnitude of input parameter  $b_n$  for the negative group increases, adding to the group’s negative opinion. It can be noted that the positive group’s input is higher than the negative group’s input.

Figure 5 presents the group opinion dynamics after applying the optimal parameter settings obtained from the calibration process and updating agents’ opinions as per equations (1) and (2) accordingly. This figure shows how the model, fine-tuned with these parameters, captures shifts in group opinions over time, reflecting a more realistic and potentially accurate representation of trader’s group opinions compared to the baseline. The dynamics here consider interactions and feedback mechanisms that the optimal parameters enable. Since the negative group has the lower self-reinforcement and the lower inputs, I found that the best instances of the model’s evolution happen when the positive group update their opinion at a higher rate, which means as the initial opinions are  $-1$  and  $+1$ , for the negative group and the positive group respectfully, the negative group get to keep its negative average opinion close to the initial opinion, and the dynamics in (1) will results in negative opinion. The jump in opinions happens around days 500 or 900 when the positive group becomes more self-reinforcing ( $\alpha_p > \alpha_n$ ) and ( $\alpha_p > 0.9$ ); and receives more inputs ( $b_p > b_n$ ). Thus,  $x_p$  becomes greater and updating (2) will give a positive opinion. It is noteworthy to mention that since the prediction market expects opinions to be in  $[-1, +1]$ , opinions need to be scaled. I use this simple linear transformation that maps the values of  $\frac{x_i+1}{2}$ .

Figure 6 shows that the market prices resulting from using the calibrated model capture the general

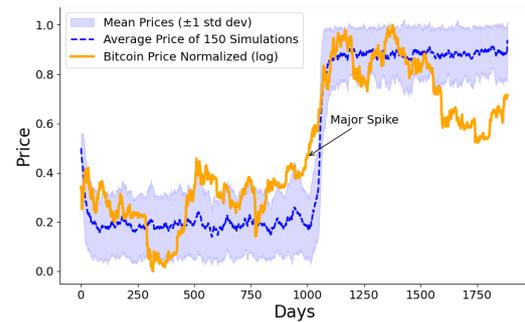


Figure 6: This figure shows the simulated average price from 150 model simulations (blue line) with the actual, normalized Bitcoin price (orange line) over the same time period.

upward and downward trends in Bitcoin’s price but lack some of the high-frequency volatility present in the actual Bitcoin market. The model provides a good approximation of the overall trend but may not fully capture the rapid fluctuations or extreme volatility seen in real-world cryptocurrency markets. These results demonstrate the model’s capacity to follow long-term price movements while highlighting potential areas for improving its responsiveness to sudden market shifts.

The distribution of residuals represents the differences between Bitcoin prices and the simulated prices generated by the calibrated model. These residuals are centered around zero, indicating that the model is not biased. However, the residuals are not perfectly normally distributed, with a mean of  $\mu = 0.00$  and a standard deviation of  $\sigma = 0.16$ . The skewed spread and shape of the residuals suggest that the model may not fully capture the underlying distribution of the data.

To evaluate the relationship between simulated prices and Bitcoin prices, Exponential Moving Averages (EMAs) and residual analysis are applied. First, EMAs are calculated for both simulated prices and Bitcoin prices to smooth fluctuations caused by the high frequency of trading in the simulation. EMAs are optimized through a minimization process aimed at reducing the variance of the residuals. After calculating the residuals, zero-crossings—points where the residuals change sign—are identified. These zero-crossings segment the time series into distinct phases, each representing periods of consistent residual behavior. The mean residual for each phase is calculated to summarize the overall behavior during that phase. Linear regression is applied to the residuals within each phase to examine trends. Figure 7 (top) shows the EMAs for both series, and Figure 7 (bottom) displays the phase means during specific time phases alongside the EMA of residuals time series.

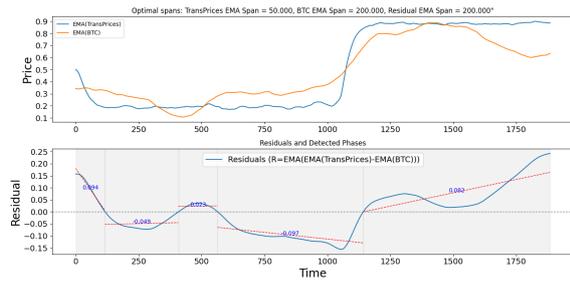


Figure 7: Comparison of Exponential Moving Averages (EMAs) of simulated prices and Bitcoin (BTC) prices. The top plot shows the EMAs of simulated prices (blue) and Bitcoin prices (orange) over time, with optimal span values of 50 days for simulated prices and 200 days for Bitcoin prices. The bottom plot displays the residuals, smoothed using an EMA with a span of 200 days, calculated as the difference between the EMAs of simulated prices and Bitcoin prices. The red dashed lines represent the linear regression of the residuals, with corresponding means annotated.

In Phase 1 ( $t_0$  to  $t_{116}$ ), the mean of residuals is positive (0.094), indicating that simulated prices are higher than Bitcoin prices. During Phase 2 ( $t_{116}$  to  $t_{408}$ ), the mean of residuals turns negative (-0.049), showing that Bitcoin prices are higher. In Phase 3 ( $t_{408}$  to  $t_{563}$ ), the mean of residuals is slightly positive (0.023), with simulated prices being higher again. Phase 4 ( $t_{563}$  to  $t_{1142}$ ) is characterized by a negative mean of residuals (-0.097), indicating a pronounced divergence where Bitcoin prices are much higher. Finally, in Phase 5 ( $t_{1142}$  to  $t_{1884}$ ), the mean of residuals becomes positive once more (0.082), showing a return to simulated prices being higher. On average, the residuals are close to zero relative to the normalized price scale (0 to 1), suggesting that the model performs reasonably well. However, during certain periods, such as Phase 4, the residuals are relatively larger, indicating that the model’s predictions are less accurate.

Table 1: Comparison of Performance Metrics for the calibrated model and the baseline model.

Metric	Calibrated Model	Baseline
MSE	0.03	0.12
MAE	0.13	0.29
RMSE	0.16	0.35
R <sup>2</sup>	0.77	-0.85
Pearson Correlation	0.88	0.23

Table 1 provides a comparison of the calibrated model and the baseline model across several performance metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R-squared (R<sup>2</sup>), and Pearson Correlation.

The calibrated model shows a significant improvement over the baseline in all metrics. With an MSE of 0.026 compared to 0.124 for the baseline, the calibrated model demonstrates more accurate predictions with smaller average squared deviations from the true values. Similarly, the MAE is much lower (0.133 for the calibrated model vs. 0.288 for the baseline), indicating that the predictions are generally closer to actual values. The RMSE, which gives more weight to larger errors, is also considerably lower for the calibrated model (0.161 versus 0.352), highlighting fewer large deviations from the actual data.

Additionally, the calibrated model achieves an R-squared value of 0.772, meaning it explains approximately 77.2% of the variance in the data, while the baseline model has a negative R-squared, indicating a poor fit. Lastly, the Pearson Correlation for the calibrated model is 0.883, showing a strong positive linear relationship between the predicted and actual values, compared to a weaker correlation of 0.230 for the baseline. Overall, these metrics underscore the superior accuracy and predictive capability of the calibrated model over the baseline.

### 3.3 Financial Market Forecasting

We can use the calibrated parameters from the model of the prediction market to forecast the financial market. We can formulate the problem of short-term forecasting of asset prices and group opinion dynamics as *multi-target regression* (MTR) problem. Formally, given an input vector  $\mathbf{X} \in \mathbb{R}^n$  representing the features—parameter values, market mood and price history—, the goal is to predict an output vector  $\mathbf{Y} \in \mathbb{R}^m$  where each element  $y_i$  represents the predicted value for a specific target variable (e.g., future prices at different time horizons), and the expected group dynamics. This formulation allows us to leverage *supervised machine learning* techniques to learn a mapping function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  such that:

$$\mathbf{Y} = f(\mathbf{X}) + \varepsilon$$

where  $\varepsilon$  denotes a vector of error terms, capturing the model’s prediction error. MTR problems can be solved by a variety of machine-learning algorithms. Given that the generated dataset contains targets that are both continuous and categorical, interdependent, and the dataset size is small. It is important to use multi-target supervised machine learning methods that are capable of effectively handling target interdependencies while at the same time avoiding overfitting due to limited data.

Table 2 presents the performance metrics for several multi-target learning algorithms, Multi-Output

Random Forest (MORF), a variant of Random Forest, models interdependencies within tree splits and is particularly robust for small datasets, making it an effective choice for multi-target tasks. Extreme Gradient Boosting Trees (XGBoost) support multi-target prediction and perform exceptionally well on smaller datasets. Finally, Multi-Task Learning Neural Networks (MTNN) leverage shared hidden layers to capture features common across all targets while utilizing separate output layers tailored to handle continuous and categorical variables. MTXGBoost generally achieves the best performance across most targets, followed closely by MORF, while MTNN shows slightly higher errors but remains competitive in certain cases.

Table 2: Performance Metrics for Multi-target Models.

Model	Target	MAE	RMSE	R <sup>2</sup>
MTXGBoost	Price	0.16	0.24	0.86
	$\alpha_p$	0.02	0.04	0.82
	$\alpha_n$	0.03	0.05	0.77
	$b_p$	0.11	0.16	0.84
	$b_n$	0.11	0.16	0.83
MORF	Price	0.17	0.25	0.85
	$\alpha_p$	0.02	0.04	0.80
	$\alpha_n$	0.03	0.05	0.76
	$b_p$	0.12	0.16	0.83
	$b_n$	0.12	0.17	0.82
MTNN	Price	0.16	0.24	0.86
	$\alpha_p$	0.03	0.04	0.76
	$\alpha_n$	0.04	0.05	0.70
	$b_p$	0.12	0.16	0.83
	$b_n$	0.13	0.17	0.82

## 4 CONCLUSION

In this paper, I explored price formation in prediction markets, by populating a simple model of a prediction market exchange with opinionated agents and calibrating model parameters. I investigated the influence of key drivers in the opinion formation process within two competing groups: self-reinforcement, herding, and additive responses to inputs. Using a real-world dataset, the results indicate that both groups generally follow overall market sentiment. However, when market mood aligns with a particular group’s opinion, that group becomes more self-reinforcing, while a lack of alignment reduces self-reinforcement.

Despite these insights, there are limitations to this simple ABM. The model’s simplified structure does not fully capture the complexity of real-world trading behavior, particularly the high-frequency volatil-

ity and external shocks often observed in cryptocurrency markets. Additionally, while the model accurately reflects long-term sentiment trends, it may fail to capture the rapid sentiment shifts that occur in actual trading environments. Future work could address these limitations by incorporating more complex agent interactions or external factors influencing market sentiment. Nevertheless, the model introduced here does capture real-world price dynamics with sufficient accuracy to be of significant and enduring interest. To enable other researchers to replicate and extend this work I have made the source code and sample data sets available on GitHub.

## REFERENCES

- Benedetti, M., Catapano, G., Sclavis, F. D., Favorito, M., Glielmo, A., Magnanimi, D., and Muci, A. (2022). Black-it: A ready-to-use and easy-to-extend calibration kit for agent-based models. *Journal of Open Source Software*, 7(79):4622.
- Berg, J. E. and Rietz, T. A. (2003). Prediction markets as decision support systems. *Information systems frontiers*, 5:79–93.
- Bizyaeva, A., Franci, A., and Leonard, N. E. (2020). A general model of opinion dynamics with tunable sensitivity. *arXiv preprint arXiv:2009.04332*.
- Bizyaeva, A., Franci, A., and Leonard, N. E. (2022). Nonlinear opinion dynamics with tunable sensitivity. *IEEE Transactions on Automatic Control*, 68(3):1415–1430.
- Glielmo, A., Favorito, M., Chanda, D., and Delli Gatti, D. (2023). Reinforcement learning for combining search methods in the calibration of economic abms. In *Proceedings of the Fourth ACM International Conference on AI in Finance*, pages 305–313.
- Leonard, N. E., Keena Lipsitz, Anastasia Bizyaeva, Alessio Franci, and Yphtach Lelkes (2021). The nonlinear feedback dynamics of asymmetric political polarization. *Proceedings of the National Academy of Sciences*, 118(50):e2102149118.
- Luckner, S., Schröder, J., Slamka, C., Skiera, B., Spann, M., Weinhardt, C., Geyer-Schulz, A., and Franke, M. (2011). *Prediction markets: Fundamentals, designs, and applications*. Springer Science & Business Media.
- Restocchi, V., McGroarty, F., Gerding, E., and Brede, M. (2023). Opinion dynamics explain price formation in prediction markets. *Entropy*, 25(8):1152.
- Song, B., Xiong, G., Yu, S., Ye, P., Dong, X., and Lv, Y. (2021). Calibration of agent-based model using reinforcement learning. In *2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)*, pages 278–281. IEEE.